Block replication 副本数

dfs.blocksize block大小

# blocksize设置

现场一个block大小128M。测试现场硬盘写入速度及机器间网速如下：







HDFS中平均寻址时间大概为10ms；经过官方测试发现，寻址时间为传输时间的1%时，为最佳状态；所以最佳传输时间为10ms/0.01=1000ms=1s。

目前测试磁盘的传输速率>190MB/s , 机器间传输速率只有95MB/s；

计算出最佳block大小：95MB/s x 1s = 95MB，所以设定block大小为128MB是合理的。

**方案1：提升网速至少为200M以上，则可配置block size为256M，可有效降低内存占用。**

# 修改内存配置

现场环境显示文件\文件夹和bolck对象共有165866260个，参照如下对照表

可将NameNode Java heap size修改为92160M-->94976M，NameNode new generation size和NameNode maximum new generation size修改为11520M-->8960M



**方案2：修改配置NameNode Java heap size：94976M，NameNode new generation size：8960M，NameNode maximum new generation size：8960M**

# 合并小文件

目录/文件和block均会占用namenode内存，大量小文件会降低内存使用效率，小文件的读写性远远低于大文件的读写，对小文件读写需要在多个数据源切换，严重影响性能。

使用hadoop自带archive对集群中的小文件进行归档，将小文件打包到更大的HAR文件。一般建议HAR文件和blocksize大小接近。

* **用法：**



-archiveName：指定归档文件名称

-p：指定要归档目录或文件的父目录

-r：指定归档文件的副本数

\*：使用通配的方式归档一类文件或目录

dest：指定归档文件的输出目录

* **举例：**

**待合并小文件**

****

**合并：**hadoop archive -archiveName test.har -p /zc/small/small -r 2 \* /zc/small



当创建archive时原文件不会删除，需手动删，创建是由mapreduce作业完成的。



**查看归档文件和原文件一样**

****

**其他可能用到的操作命令**

将归档文件内容复制到其他目录 hdfs dfs -cp har:///zc/small/test.har/\* /zc/

删除归档文件 hdfs dfs -rm -r /zc/small/test.har

# 定期删除没用的文件